

Parametrization of the GMIPp for the Study of Stacking Interactions

BEGOÑA HERNÁNDEZ,¹ F. JAVIER LUQUE,² MODESTO OROZCO¹

¹*Departament de Bioquímica i Biologia Molecular, Facultat de Química, Universitat de Barcelona, Martí i Franquès 1, Barcelona 08028, Spain*

²*Departament de Físicoquímica, Facultat de Farmàcia, Universitat de Barcelona, Avda Diagonal s/n, Barcelona 08028, Spain*

Received 7 December 1998; accepted 8 February 1999

ABSTRACT: A parametrization of the Generalized Molecular Interaction Potential with polarization (GMIPp) for the study of aromatic stacking is presented. Parametrization is limited to the van der Waals parameters defining the dispersion–repulsion interactions between the quantum and classical molecules, while electrostatic and polarization contributions are computed at the quantum mechanical level using a perturbational approach. The new parametrized GMIPp has been used to study different stacking interactions between nucleic acid bases. © 1999 John Wiley & Sons, Inc. *J Comput Chem* 20: 937–946, 1999

Keywords: GMIPp; stacking interactions; parametrization

Introduction

The quantum mechanical (QM) study of molecular interactions is typically performed using the “supermolecule” approach. In this approach the interaction energy is determined from the difference between the energy of the complex and the energy of the isolated, fully relaxed

monomers. The interaction energy estimated in this way is overestimated, owing to the neglect of the basis set superposition error,¹ which is generally corrected using the Boys–Bernardi procedure.² This implies that for a dimerization process four additional calculations have to be performed to obtain the corrected interaction energy. Supermolecule approach can yield accurate estimates of dimerization energy if high levels of theory are used, but at a very large computational cost.

The computational cost of the calculation of dimerization energy is sensibly reduced if the QM treatment is limited to one of the monomers (A), while the other (B) is described classically. In this case the interaction energy is determined from eq.

Correspondence to: M. Orozco; e-mail: modesto@luz.bq.ub.es

Contract/grant sponsor: Spanish DGICYT; contract/grant numbers: PB96-1005 and PB97-0908.

Contract/grant sponsor: Centre de Supercomputació de Catalunya.

(1), where Ψ^A is the wave function of the isolated monomer A, and Ψ^{AB} denotes the wave function of A in presence of B. This latter wave function is typically obtained by solving the Schrödinger equation [eq. (2)] for the effective Hamiltonian (\hat{H}^{ef}) given in eq. (3) using the standard SCF procedure.

It should be noted that mixed QM/MM formalism assumes: (a) that dispersion–repulsion interactions can be reproduced by an empirical term; (b) that charge transfer between QM and classical monomers is negligible, and (c) that polarization of the classical particle by the QM monomer can be accounted for in the force-field parametrization. Even though the mixed QM/MM approach is clearly less rigorous than the pure QM, the lower expansiveness of the calculation explains their increasing application to the study of chemical systems in condensed phases and biochemical systems.^{3, 4}

$$E = \langle \Psi^{AB} | \hat{H}^{\text{ef}} | \Psi^{AB} \rangle - \langle \Psi^A | \hat{H}^0 | \Psi^A \rangle \quad (1)$$

$$\hat{H}^{\text{ef}} \Psi^{AB} = E^{AB} \Psi^{AB} \quad (2)$$

$$\hat{H}^{\text{ef}} = \hat{H}^0 + \sum_{i=1}^N \frac{Q_i}{r} + E_{vW} \quad (3)$$

where H^0 is the Hamiltonian of the isolated monomer A, Q_i denotes the point charges that represent the charge distribution of monomer B, and E_{vW} stands for the classical Lenard–Jones dispersion–repulsion interaction between A and B.

Another computational alternative to determine the interaction energy relies on the use of the perturbation methods.⁵ Basically, the dimer is conceived as a perturbation of the corresponding non-interacting monomers, and the energy contributions are determined from different perturbative expressions. This approach can easily be translated to mixed simulations, where it benefits from the fact that only the wave function of the isolated QM monomer needs to be determined in a previous calculation. Following the same treatment noted above, the QM/MM interaction energy can be expressed by adding electrostatic, polarization, and dispersion–repulsion terms. In this context we have recently proposed a QM/MM method,^{6–10} named Generalized Molecular Interaction Potential with polarization GMIPp [eq. (4)].

The GMIPp is derived as a generalization of the molecular electrostatic potential (MEP),¹¹ and allows the calculation of the interaction energy between a QM particle and any polyatomic charge distribution, including a perturbational correction

for classical–quantum polarization effects,^{9, 10, 12} and an empirical dispersion–repulsion term (the van der Waals term). The GMIPp provides QM–classical interaction energies comparable to the values derived from a “supermolecule” approach, but avoiding the need to recompute the wave function of the QM monomer perturbed by the classical particle.⁹ This implies a notable reduction of the computational cost, especially in cases where many configurations of the classical–QM system are computed.

$$\begin{aligned} \text{GMIPp} = & \sum_A \sum_B \frac{Q_B Z_A}{|R_B - R_A|} \\ & - \sum_B Q_B \sum_i^{\text{occ}} \sum_\mu \sum_\nu c_{\mu i} c_{\nu i} \left\langle \phi_\mu \left| \frac{1}{|R_B - r|} \right| \phi_\nu \right\rangle \\ & + \sum_{A'B'} \left(\frac{C_{A'B'}}{|R_{B'} - R_{A'}|^{12}} - \frac{D_{A'B'}}{|R_{B'} - R_{A'}|^6} \right) \\ & + \sum_j^{\text{vir}} \sum_i^{\text{occ}} \frac{1}{\varepsilon_i - \varepsilon_j} \\ & \times \left\langle \sum_\mu \sum_\nu c_{\mu i} c_{\nu j} \left\langle \phi_\mu \left| \sum_B \frac{Q_B}{|R_B - r|} \right| \phi_\nu \right\rangle \right\rangle^2 \end{aligned} \quad (4)$$

where c and ε stand for the coefficients and energies of the molecular orbitals defining the unperturbed wave function of the QM particle, C and D are empirical van der Waals parameters, and the subindexes A' and B' denote the centers in QM and classical particles having van der Waals parameters.

The GMIPp has been used to study different chemical interactions,^{8–10, 13} and it is expected to be of general applicability to a wide variety of systems. A particularly interesting interaction is π -stacking of aromatic compounds. This interaction occurs frequently in molecular recognition, and is also essential in protein stabilization,¹⁴ host–guest recognition,^{15, 16} and in the mechanism of the action of several drugs.¹⁷ Furthermore, stacking interactions contributes decisively to the structure and reactivity of nucleic acids.^{16, 18, 19} The study of stacking interactions from pure QM methods is extremely difficult. First, because these interactions are dominated by a combination of electrostatic and dispersion effects,^{20–23} very large basis sets and high levels of theory to include correlation effects are required in QM calculations.^{20–24} Second, because stacking interac-

tions exhibit a reduced directionality, the configurational space of stacked dimers is complex and needs to be explored systematically, which implies that perhaps hundreds of high-level QM calculations are necessary to reproduce the configurational space of the stacked dimer. It is clear that such a kind of calculation is not yet affordable with current computational resources.

In this article we report a specific parametrization of the GMIPp for the study of stacking interactions. The GMIPp is parametrized against very high-level QM calculations of model systems, and is subsequently tested in a variety of stacking interactions in different nucleic acid systems.

Methods

DETAILS OF GMIPp CALCULATIONS

The van der Waals contribution is the only term that typically has to be parametrized in GMIPp calculations. Parametrization can be done by fitting interaction energies determined at the “supermolecule” level with GMIPp estimates obtained with the same wave function. However, there are cases where it is advisable to fit the GMIPp computed from an inexpensive wave function to reference data determined experimentally or, from very high-level QM calculations. This approach seems

advisable for instance in cases where dispersion effects are expected to be important, arguing against the use of simple SCF calculations as a reference. In these cases, an extraparametrization effort might be necessary to account for the different magnitude of electrostatic and polarization terms in the high-level QM (or experiment) and GMIPp calculation.

In this study, following our previous works, we decided to determine the GMIPp from HF/6-31G(d) wave functions, while the reference data were obtained from CCSD(T)/cc-pvdz (benzene dimer) or MP2/cc-pvdz calculations (the rest of dimers). Because the HF/6-31G(d) method maximizes the MEP₂₅ with respect to high-level estimates, we might expect an overestimation of strength of electrostatic interactions in HF/6-31G(d) calculations with regard to the reference calculations. To correct this problem [intrinsic to HF/6-31G(d) calculations], and to guarantee the correct balance between electrostatic and nonelectrostatic contributions, the electrostatic term in the GMIPp was scaled [see parameter α in eq. (5)]. Such a scaling factor was determined (prior to any fitting of the van der Waals term) by fitting the HF/6-31G(d) and MP2/6-31++G(d,p) QM MEPs computed for 97 configurations of the dimers of benzene, imidazole, pyrimidine, and 2-amino,4-oxypyrimidine (see Fig. 1). The optimized scaling

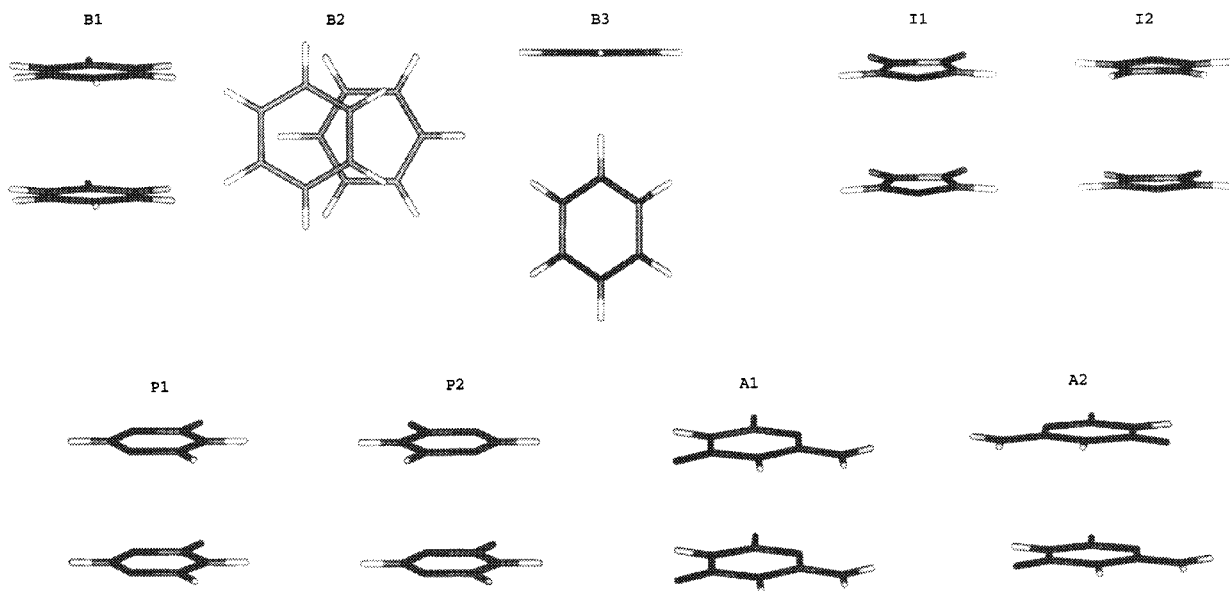


FIGURE 1. Dimers considered in the parametrization of the GMIPp.

factor was 0.888, which *a priori* agrees with the expected overestimation of polarity in HF/6-31G(d) calculations (ref. 25, and references therein). This factor was used in all the following calculations. No scaling of the polarization term is necessary; even we can expect a certain underestimation of the polarization energy in HF/6-31G(d) calculations.⁹

$$\begin{aligned} \text{GMIPp} = & \alpha \sum_A \sum_B \frac{Q_B Z_A}{|R_B - R_A|} \\ & - \alpha \sum_B Q_B \sum_i^{\text{occ}} \sum_{\mu} \sum_v c_{\mu i} c_{v i} \\ & \times \left\langle \phi_{\mu} \left| \frac{1}{|R_B - r|} \right| \phi_v \right\rangle \\ & + \sum_{A'B'} \left(\frac{C_{A'B'}}{|R_{B'} - R_{A'}|^{12}} - \frac{D_{A'B'}}{|R_{B'} - R_{A'}|^6} \right) \\ & + \sum_j^{\text{vir}} \sum_i^{\text{occ}} \frac{1}{\epsilon_i - \epsilon_j} \\ & \times \left\langle \sum_{\mu} \sum_v c_{\mu i} c_{v j} \left\langle \phi_{\mu} \left| \sum_B \frac{Q_B}{|R_B - r|} \right| \phi_v \right\rangle \right\rangle^2 \end{aligned} \quad (5)$$

The van der Waals term was parametrized by minimizing the error function [see eq. (6)] for those stacked configurations of the dimers (see Fig. 1). Atomic charges for the classical molecules were determined by fitting the MP2/6-31++G(d) MEPs to an atom-centered Coulombic expansion. Standard combination rules were used in the van der Waals expression [see eq. (7)]. The van der Waals parameters of the classical particle were derived from OPLS values^{26,27}

$$f(vW) = |E_{\text{QM}} - \text{GMIPp}| \quad (6)$$

where E_{QM} is determined considering the interacting monomers at a high QM level, and the GMIPp is computed as noted in eq. (5).

$$\begin{aligned} E_{vW} = & \sum_{A'B'} (\epsilon_{A'} \epsilon_{B'})^{1/2} \\ & \times \left(\frac{(R_{A'}^* + R_{B'}^*)^{12}}{|R_{B'} - R_{A'}|^{12}} - 2 \frac{(R_{A'}^* + R_{B'}^*)^6}{|R_{B'} - R_{A'}|^6} \right) \end{aligned} \quad (7)$$

DEFINITION OF THE QM-CLASSICAL PARTITION

QM/MM methods are not invariant to permutation between classical and QM particles, that is, the interaction energy can be different if the QM monomer is treated classically and vice versa. This problem arises from the uncertainties in the set of point charges assigned to the classical particle, and eventually to the different polarizability of the interacting monomers (note that only the classical-QM polarization is treated explicitly in QM/MM calculations). In the present study, the polarization contribution was rather small, typically less than 0.2 kcal/mol, and accordingly, the latter aspect was not relevant for the above-mentioned problem. To correct the underlying uncertainty, all electrostatic calculations [first term in eq. (5)] were repeated, interchanging the treatment (QM and classical) of the two interacting monomers. The final electrostatic energy was determined as the average of the two corresponding values.

COMPUTATIONAL DETAILS

Experimental (if available) or HF/6-31G(d,p) optimized geometries²⁸ of planar monomers were used in calculations. The geometries of the monomers were frozen in the different dimer configurations. The wave functions and QM energies were determined using a Gaussian-94 computer program.²⁹ The MEP and GMIPp were computed using a MOPETE³⁰ computer program, and when necessary MOPFIT³¹ was used to derive ESP charges. All calculations were performed on the Origin-2000 of the Centre de Supercomputació de Catalunya (CESCA), and on workstations in our laboratory.

Results and Discussion

GMIPp PARAMETRIZATION

The van der Waals parametrization was accomplished considering only five atom types for the molecules considered in the study: carbon, oxygen, nitrogen (unbound to hydrogen), nitrogen (bound to hydrogen), and hydrogen bound to carbon atoms. Following the OPLS force field,^{26,27} hydrogens bound to heteroatoms were not considered in the parametrization. The optimized parameters are shown in Table I. Comparison of the optimized

TABLE I.
Optimized van der Waals Parameters (ϵ , in kcal/mol;
 R^* , in Å) for the QM Atoms.

Atom	ϵ (QM)	R^* (QM)	ϵ (OPLS)	R^* (OPLS)
C	0.0233	2.1592	0.07 ^a	1.99 ^a
C			0.08	1.96
O	0.0543	1.6451	0.21	1.66
N	0.0500	1.7538	0.17	1.82
N(H)	0.0288	1.9124	0.17	1.82
H(C)	0.0167	1.1608	0.05	1.40

The OPLS-adapted classical van der Waals parameters are given for comparison.

^aOPLS van der Waals parameters for the carbon atom of benzene and heterocycles are slightly different.

values with the OPLS ones indicates a reasonable similarity, which contrast with previous studies where we found that QM/MM and purely classical van der Waals parameters used to represent hydrogen bonded structures are quite different.⁸ We should note here that optimum QM/MM van der Waals parameters presented in Table I can be used with confidence to represent stacking interactions, such as hydrogen bonds, where we might expect larger and “softer” QM nuclei.⁸ For instance, the “stacking-optimized” GMIPp values for the A·T and G·C Watson–Crick H-bonding are around -16 to -29 kcal/mol. These values are quite reasonable (probably within the range of accuracy of most empirical force fields), but present a larger range of error [around 4 kcal/mol with respect to B3LYP/6-31G(d) calculations] than that expected for stacked complexes.

GMIPp calculations incorporating the optimized van der Waals parameters quite accurately reproduce the interaction energies computed at

the CCSD(T)/cc-pvdz and MP2/cc-pvdz levels [CCSD(T) calculations for the benzene dimer were taken directly from a recent work by Hobza and coworkers; see ref. 21]. Thus, the correlation coefficient of the regression equation shown in Figure 2 is 0.92, and the ratio QM/GMIPp (1.05) is very close to the optimum value of 1.0, indicating the lack of systematic deviations in the GMIPp estimates. Finally, the root-mean-square deviation (rms) is only 0.4 kcal/mol, while the interaction energies vary from -3 to $+2$ kcal/mol. In summary, the GMIPp provides very fast estimates of stacking energies for the series studied with good accuracy.

The reliability of GMIPp results is also clear from comparison of the dimerization energies given in Table II. Thus, the rms deviation between

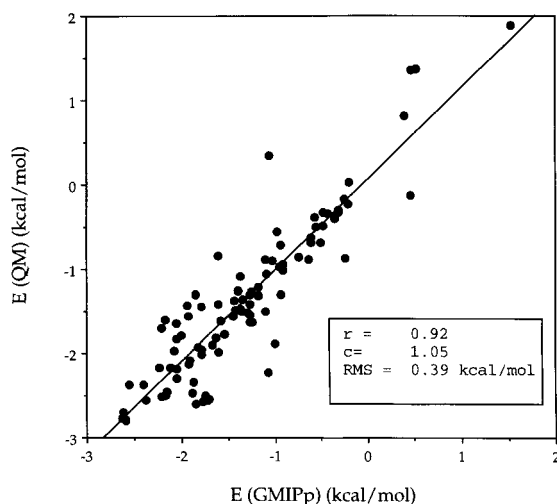


FIGURE 2. Relationship between GMIPp and high level estimates of dimerization energy computed for different configurations of the dimers displayed in Figure 1.

TABLE II.
QM and GMIPp Dimerization Energies (kcal/mol) at the Optimized Distance (Å) for Different Dimers.

Dimer	d (Å)	E (QM)	E (GMIPp)	E (ele)	E (vw)	E (pol)
I2	3.7	-2.6	-2.5	-0.9	-1.5	-0.15
B1	4.1	-1.0	-0.8	+1.1	-1.8	-0.04
B2	5.1	-2.1	-1.7	-0.4	-1.3	-0.01
B3	4.0	-1.8	-1.1	+0.8	-1.8	-0.09
P1	3.9	-1.2	-1.0	+1.1	-2.1	-0.04
P2	3.6	-2.7	-3.2	-1.1	-2.2	-0.06
A1	4.0	+0.3	+0.1	+2.7	-2.5	-0.05
A2	3.6	-3.9	-3.8	-1.1	-2.6	-0.09

The GMIPp interaction energy is decomposed into electrostatic (ele), van der Waals (vw), and polarization (pol) contributions. See Figure 1 for nomenclature.

QM and GMIPp values is 0.3 kcal/mol. Indeed, analysis of the different contributions to the GMIPp indicates that stacking interactions are typically dominated by van der Waals interactions, which are always stabilizing, as suggested previously.^{20–24} The electrostatic contribution is typically smaller, and often unfavorable to dimerization. Finally, despite the large polarizability of aromatic systems, the polarization contribution is very small and can be neglected in most cases. It is worth noting that the electrostatic component varies greatly, depending not only on the nature of the monomers, but also on their relative orientation in the dimer. On the contrary, the van der Waals term is quite invariant to changes in orien-

tation of the monomers. Therefore, while the van der Waals term are responsible for the stacking, the optimum orientation of the monomers in the dimer is dictated by the electrostatic component.

The relatively small importance of polarization effects in stacking interactions support the use of classical methods to study stacking interactions at the qualitative level, as suggested by other authors.³² However, it should be noted that more accurate results are obtained with the GMIPp compared with those obtained using empirical force fields. This is clear for in Figure 3, where several interaction profiles determined at the GMIPp and classical levels (using the same classical force-field parameters used in GMIPp calculations to describe

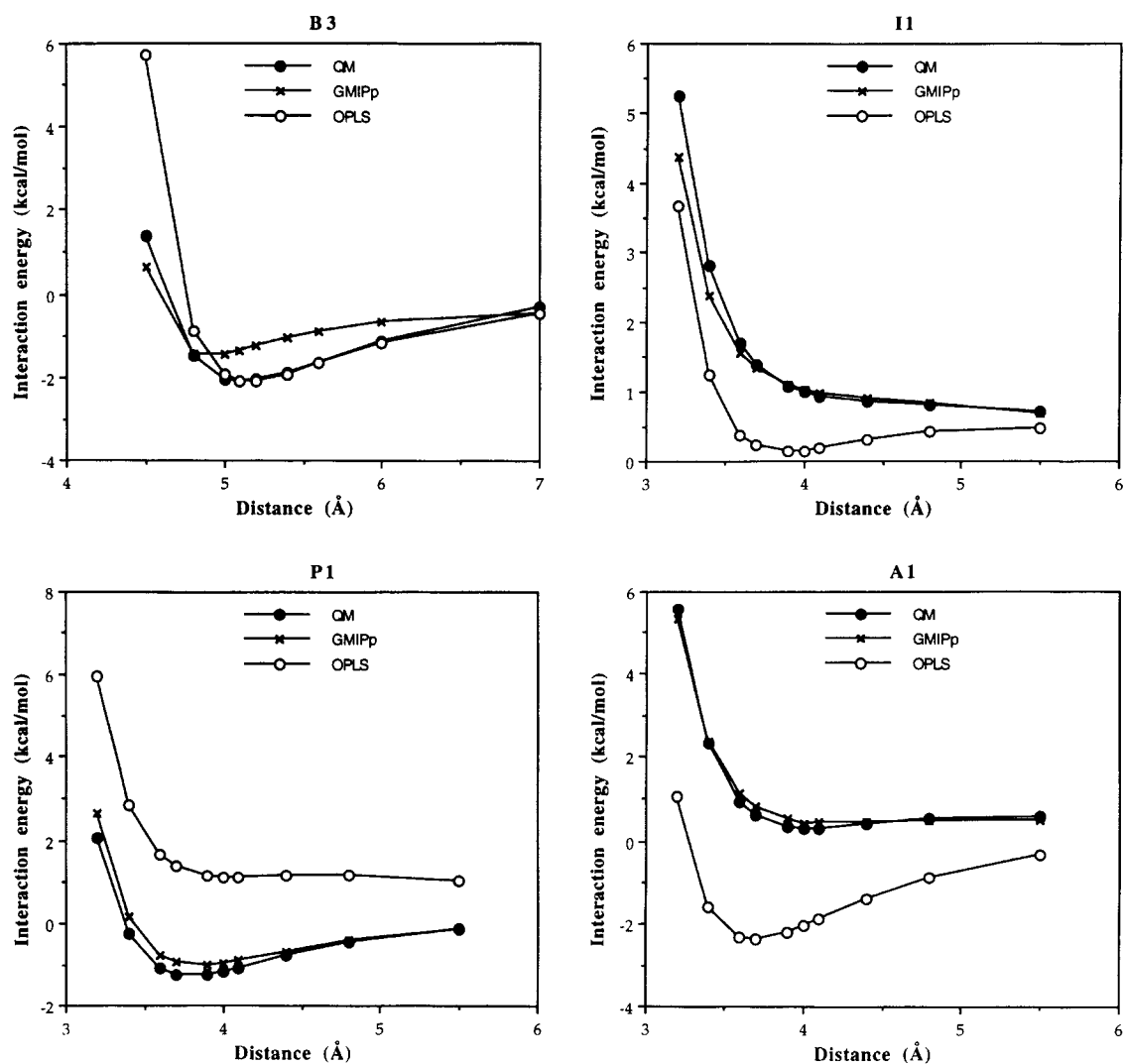


FIGURE 3. Comparison of GMIPp and classical energy profiles with high level QM data for selected dimers (see Fig. 1).

the classical particle, typically OPLS, or ESP/OPLS) are compared with those derived using high levels of QM theory.

B-TYPE STACKING OF NUCLEIC ACID BASES

To further test the suitability of GMIPp for the study of aromatic stacking, the intrastrand stacking energies of nucleic acid bases in B-DNA were determined and compared with our previously published MP2/6-31G(d) results.²⁰ Note that these QM results should not be considered as a high-level reference, but just as a rough estimate of the stacking energy. In fact, the quality of MP2/6-31G(d) estimates of stacking energies has been the subject of a large debate (refs. 20, 21, and 24, and references therein), and it seems that they might underestimate the strength of the interaction. The geometries were determined as noted elsewhere,²⁰ following Arnott's helical parameters for B-type DNA.³³

Results in Table III and Figure 4 show the good ability of the GMIPp to reproduce the MP2/6-31G(d) values both in relative and absolute terms. Thus, the rms deviation is only 1.0 kcal/mol, clearly below the expected limit of accuracy of MP2/6-31G(d) calculations. It is worth noting that the quality of the HF estimate is worse than the much less expensive GMIPp calculation. Figure 4 also shows that the interchange between QM and

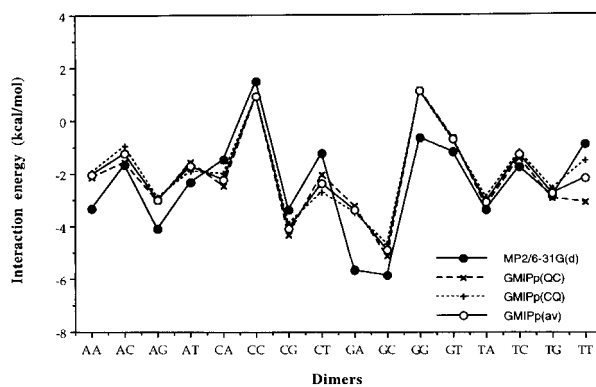


FIGURE 4. Interaction energies for the different nucleic acid base pairs stacked in the B-type conformation computed using MP2/6-31G(d) and GMIPp methods.

classical particles in GMIPp calculations leads to negligible changes in interaction energies. Overall, the results support the suitability of the GMIPp to study stacking interactions in nucleobases instead of the most costly HF or MP2 calculations.

Inspection of results in Table III points out that, as noted elsewhere,^{20, 22, 34–40} dispersion interactions (van der Waals) are extremely important for stacking. In fact, in most cases van der Waals is the leading term, favoring stacking, while electrostatic is in general smaller (in absolute values), and in 50% of the cases disfavors stacking. The van der Waals stabilizing effect is generally larger for

TABLE III.
MP2/6-31(d) and GMIPp Estimates of the Intrastrand Stacking Energy (kcal/mol) for Nucleic Acid Base Dimers in the B-type Conformation.

Dimer	$E(\text{ele})$	$E(\text{pol})$	$E(\text{vW})$	$E(\text{GMIPp})$	$E^{\text{MP2/6-31G(d)}}$
AA	0.57	-0.11	-2.51	-2.06	-3.34
AC	0.99	-0.22	-2.03	-1.26	-1.68
AG	-0.03	-0.27	-2.68	-2.98	-4.10
AT	0.41	-0.17	-1.97	-1.73	-2.31
CA	0.28	-0.28	-2.24	-2.24	-1.48
CC	2.91	-0.28	-1.74	0.89	1.46
CG	-1.44	-0.40	-2.27	-4.11	-3.39
CT	-0.31	-0.31	-1.78	-2.40	-1.26
GA	-0.92	-0.15	-2.30	-3.37	-5.67
GC	-2.70	-0.24	-1.97	-4.90	-5.86
GG	3.87	-0.27	-2.49	1.12	-0.67
GT	1.54	-0.19	-2.07	-0.72	-1.19
TA	-0.69	-0.13	-2.26	-3.07	-3.37
TC	0.62	-0.15	-1.78	-1.31	-1.78
TG	-0.27	-0.24	-2.27	-2.78	-2.86
TT	-0.22	-0.20	-1.73	-2.17	-0.90

The GMIPp values, which are decomposed into electrostatic, polarization, and van der Waals components, are the average of two calculations where QM and classical particles were interchanged. The MP2/6-31G(d) values are from ref. 20.

purine · purine stackings than for pyrimidine · pyrimidine ones, as expected from the greater surface overlap in the former. However, the van der Waals interaction lies in a close range (from -1.7 to -2.7 kcal/mol) within the series, and most of the difference in stability between the pairs comes from the electrostatic term, which varies from 3.9 to -2.7 kcal/mol. The polarization contribution is small (around -0.2 kcal/mol in average), but not negligible. In summary, as expected from previous pure QM calculations,^{20,22,34–40} the strength of stacking is given by dispersion interactions, while the selectivity of such interactions is provided by electrostatic interactions.

GEOMETRY-DEPENDENT CHANGES IN STACKING

On the basis of the preceding results, we extended our study to examine how changes in the internal geometry of the dimer influence the stacking energy in the gas phase. This type of analysis can be useful to determine the intrinsic characteristics of stacking, which combined with the solvent effect on these systems^{41,42} defines the stacking in physiological conditions. Particularly, we explored the dependence of the stacking energy of the AC dimer on the rise (the separation distance between stacked bases) and twist (the rotation between bases) in the gas phase. In all cases the B-type geometry was adopted, but for those parameters defining rise and twist.

The energy profile obtained from different values of rise is given in Figure 5. The GMIPp profile follows the change in van der Waals energy, which is the main contribution to stacking for the interval of rise comprised between 3.0 and 5.2 Å (the elec-

trostatic term is the leading component beyond 5.2 Å). As expected, the GMIPp is repulsive for short base–base distances (rise < 3.2 Å), while is attractive even at distances of 5.2 Å (-0.22 kcal/mol), indicating that stacking is not just a pure short-term interaction. The GMIPp minimum appears at rise values around 3.6 Å, not far from the values found in physiological DNAs.^{19,33}

The electrostatic term is always positive, and is quite constant in the range of rise values studied (it decreases only 0.3 kcal/mol when the rise increases from 3 to 5.2 Å). The polarization contribution is favorable by a magnitude ranging from -0.5 kcal/mol to 0.05 kcal/mol for the interval considered. As expected, the decay with intermolecular distance of the polarization contribution is faster than that of the electrostatic term. In the GMIPp minimum, the polarization (-0.26 kcal/mol) is around 30% the magnitude of the electrostatic interaction (0.86 kcal/mol), which represents a nonnegligible contribution.

To check the reliability of the GMIPp profile, we determined the stacking (BSSE-corrected) energy at the HF/6-31G(d) and MP2/6-31G(d) levels for rise values of 3.4 , 3.6 , 3.8 , and 4.6 Å. The results (see Fig. 6) show the quality of the GMIPp values, which reproduce within few tenths of a kcal/mol the much more expensive MP2/6-31G(d) values. As expected, the HF/6-31G(d) profile is completely meaningless, owing to the lack of dispersion interactions in the calculation of the total energy.

Finally, we examined the GMIPp profile for helical twist of the AC pair (for a rise of 3.4 Å). Inspection of Figure 7 shows that there are two narrow and one wide minima. The wide minimum

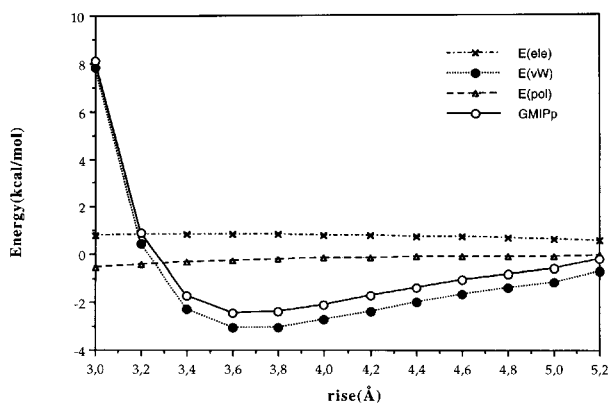


FIGURE 5. Dependence of the AC stacking energy with rise.

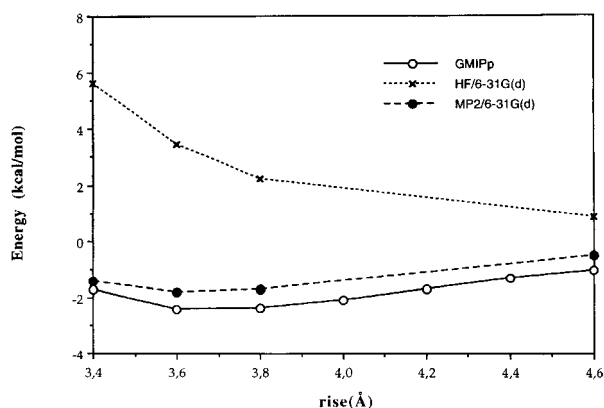


FIGURE 6. Comparison of the energy / rise profile for the AC dimer computed at the GMIP, HF, and MP2 levels.

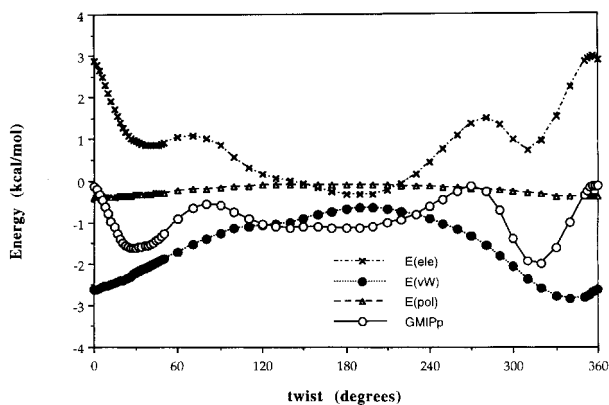


FIGURE 7. Dependence of the AC stacking energy with twist.

extends from 120 to 240 degrees, and the GMIPp minimum is around -1.1 kcal/mol. The two narrow minima are located in the dextro and levo helical regions. The dextro minimum is the deepest one (-2.6 kcal/mol), and appears at twist values around -40 degrees, in agreement with the twist value in physiological B-DNA.^{19,33} The levo minimum (-2.2 kcal/mol) is slightly wider, and appears around 30 degrees, which matches the twist value found experimentally in levo forms of DNA.^{19,33} The barriers for levo-dextro transitions are small, around 1 kcal/mol, which agrees with the well-known flexibility of DNA.¹⁹

The GMIPp profile is mainly determined by electrostatic interactions, even though the van der Waals term is typically the leading contribution. The electrostatic contribution is unfavorable for all twist values, except those corresponding to an extended conformation of DNA. The change of polarization with the twist angle follows a onefold periodicity profile and is optimum (-0.4 kcal/mol) for twist values near zero, while the less favorable contribution occurs in the extended region (-0.10 kcal/mol). Finally, the van der Waals energy also exhibits a onefold periodicity profile, with a single minimum (-2.8 kcal/mol) located for twist values around 340 degrees and a maximum (-0.7 kcal/mol) in the extended region (around 180 degrees).

Finally, Figure 8 shows the comparison of the GMIPp results with the stacking energies computed at the HF/6-31G(d) and MP2/6-31G(d) levels for selected twist values. The results demonstrate again the agreement between GMIPp and MP2/6-31G(d) profiles, and the large deviation of HF/6-31G(d) results. It is clear that GMIPp calculations can then be used to get a qualitative repre-

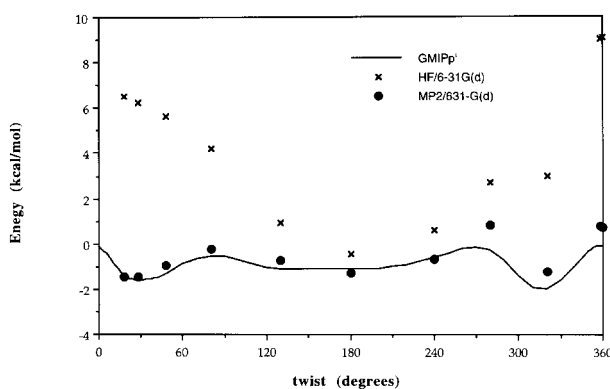


FIGURE 8. Comparison of energy / twist profile for the AC dimer computed at the GMIP, HF, and MP2 levels.

sentation of the configurational space of nucleobases stacked dimers, while much more expensive SCF calculations yield to meaningless results.

Conclusions

Parametrization of the GMIPp in front of high-level QM data allowed us to define van der Waals parameters specifically designed to be used in QM/MM studies of aromatic stacking interaction. The parametrized GMIPp is able to properly represent not only the stacked dimers considered in the parametrization process, but also several stacked dimers of nucleobases.

The GMIPp provides a very fast, simple partition of the total interaction energy, which makes possible an easy understanding of the nature of stacking. It is found that in most cases the stacking is dominated by dispersion interactions, and that the electrostatic contribution uses to be unfavorable. The magnitude of polarization effects are small, but cannot be always neglected.

Analysis of the rise and twist dependence of the AC dimer shows that while the van der Waals interaction determine the change of stacking energy with rise, the electrostatic term modulates the change of stacking interaction with twist. In all the cases the GMIPp seems to be able to reproduce not only higher level QM calculations, but also experimental data.

All the results presented in this article suggest that the GMIPp is able to provide qualitatively correct representations of the configurational space of stacked dimers at a reduced computational cost [around 50–100 times faster than equivalent (in quality) MP2 calculations]. We expect then that it

can be a method of general use in the analysis of these type of interactions.

Acknowledgments

This work has been supported by the Spanish DGICYT (PB96-1005 and PB97-0908), and by the Centre de Supercomputació de Catalunya (CESCA, Mol. Recog. Project).

References

1. Liu, B.; McLean, A. D. *J Chem Phys* 1973, 46, 3851.
2. Boys, S. F.; Bernardi, S. F. *Mol Phys* 1970, 19, 553.
3. Gao, J. *Reviews in Computational Chemistry*; Lipkowitz, K. B.; Boyd, D. B., Eds.; VCH: New York, 1996, p. 119, vol. 7.
4. Aqvist, J.; Warshel, A. *Chem Rev* 1993, 93, 2523.
5. Jeziorski, B.; Moszynski, R.; Szalewicz, K. *Chem Rev* 1994, 94, 1887.
6. Orozco, M.; Luque, F. J. In *Molecular Electrostatic Potentials: Concepts and Applications. Theoretical and Computational Chemistry*; Murray, J. S.; Sen, K., Eds.; Elsevier Science: Amsterdam, 1996, p. 181, vol. 3.
7. Orozco, M.; Luque, F. J. *J Comp Chem* 1993, 14, 587.
8. Alhambra, C.; Luque, F. J.; Orozco, M. *J Phys Chem* 1995, 99, 2084.
9. Luque, F. J.; Orozco, M. *J Comp Chem* 1998, 19, 866.
10. Cubero, E.; Luque, F. J.; Orozco, M. *Proc Natl Acad Sci USA* 1998, 95, 5976.
11. Scrocco, E.; Tomasi, J. *Top Curr Chem* 1973, 42, 95.
12. Francl, M. M. *J Phys Chem* 1985, 89, 428.
13. Colominas, C.; Orozco, M.; Luque, F. J.; Borrell, J. I.; Teixidó, J. *J Org Chem* 1998, 63, 4947.
14. Burley, S. K.; Petsko, G. A. *Adv Protein Chem* 1988, 39, 125.
15. Rebek, J. *Angew Chem Int Ed Engl* 1990, 29, 245.
16. Hunter, C. A. *Chem Soc Rev* 1994, 101.
17. Gago, F. *Methods Enzymol* 1998, 14, 277.
18. Rein, R. In *Intermolecular Interactions from Diatomics to Biopolymers*; Pullman, B., Ed.; John Wiley & Sons: Chichester, 1978, p. 307.
19. Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1988.
20. Alhambra, C.; Luque, F. J.; Gago, F.; Orozco, M. *J Phys Chem B* 1997, 101, 3846.
21. Hobza, P.; Selzle, H. L.; Schlag, E. W. *J Phys Chem* 1996, 100, 18790.
22. Sponer, J.; Leszczynski, J.; Hobza, P. *J Phys Chem* 1996, 100, 5590.
23. Hobza, P.; Selzle, H. L.; Schlag, E. W. *J Phys Chem* 1993, 97, 3937.
24. Hobza, P.; Zahradnik, R. *Chem Rev* 1988, 88, 871.
25. Soliva, R.; Orozco, M.; Luque, F. J. *J Comp Chem* 1997, 18, 980.
26. Jorgensen, W. L. *OPLS Force-Field in BOSS 3.4 Computer Program*; Yale University Press: New Haven, CT, 1990.
27. Jorgensen, W. L.; Tirado-Rives, J. *J Am Chem Soc* 1988, 100, 1657.
28. Hariharan, P. C.; Pople, J. A. *Theoret Chim Acta* 1978, 28, 213.
29. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T. A.; Peterson, G. A.; Montgomery, G. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, Y. C.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Binkley, J. S.; Defrees, D. H.; Baker, J.; Stewart, J. J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A.; *Gaussian-94 (Rev. D.3)*; Gaussian Inc.: Pittsburgh, PA, 1995.
30. Luque, F. J.; Orozco, M. *MOPETE Computer Program*; Universitat de Barcelona: Barcelona, 1998.
31. Luque, F. J.; Orozco, M. *MOPFIT Computer Program*; Universitat de Barcelona: Barcelona, 1998.
32. Hobza, P.; Kabelac, M.; Sponer, J.; Mejzlik, P.; Vondrasek, J. *J Comp Chem* 1997, 18, 1136.
33. Arnott, S.; Hukins, D. W. L. *Biochem Biophys Res Commun* 1972, 47, 1504.
34. Aida, M.; Nagata, C. *Chem Phys Lett* 1982, 86, 44.
35. Aida, M.; Nagata, C. *Int J Quant Chem* 1986, 29, 1253.
36. Hobza, P.; Sponer, J.; Polasek, M. *J Am Chem Soc* 1995, 117, 792.
37. Sponer, J.; Leszczynski, J.; Hobza, P. *J Phys Chem* 1996, 100, 1965.
38. Sponer, J.; Florian, J.; Hobza, P.; Leszczynski, J. *J Biomol Struct Dyn* 1996, 13, 827.
39. Sponer, J.; Leszczynski, J.; Vetterl, V.; Hobza, P. *J Biomol Struct Dyn* 1996, 13, 695.
40. Sponer, J.; Leszczynski, J.; Vetterl, V.; Hobza, P. *J Biomol Struct Dyn* 1996, 14, 117.
41. Friedman, R. A.; Honig, B. *Biophys J* 1995, 69, 1528.
42. Orozco, M.; Cubero, E.; Hernández, B.; López, J. M.; Luque, F. J. In *Computational Chemistry, Reviews and Current Trends*; Leszczynski, J., Ed.; World Scientific: Singapore, to appear.